the injection of exogenous RNA obtained from electrically excitable tissues into *Xenopus* oocytes causes the appearance of voltage-operated Ca^{2+} channels in the oocyte membrane. These channels are distinct from the endogenous ones in respect to their time course and inactivation properties. Moreover, heart and brain RNA each encode at least two distinct types of Ca^{2+} channels. For the heart RNA, the slow current appears to show the appropriate sensitivity to and modulation by transmitters and intracellular messengers. Significantly, a rapidly inactivating, norepinephrine- and dihydropyridine-insensitive Ca^{2+} current has recently been reported in heart cells (*10*).

The injection of RNA from various tissues into *Xenopus* oocytes should aid in the characterization of different types of Ca^{2+} channels. Such data will complement experiments ... peripheral neurons, cell lines, and reconstituted membranes. If specific RNA populations are used, it may also be possible to clarify the relations among components of Ca^{2+} channels and the interactions between channels and modulatory elements.

REFERENCES AND NOTES

1. H. Reuter, *Nature (London)* 301, 569 (1983); R. W. Tsien, *Annu. Rev. Physiol.* 45, 341 (1983).
2. M. C. Nowycky *et al.*, *Nature (London)* 316, 440 (1985); C. M. Armstrong and D. R. Matteson, *Science* 227, 65 (1985); A. P. Fox and S. Krasne, *J. Physiol. (London)* 356, 491; S. A. Fedulova *et al.*, *ibid.* 359, 431 (1985).
3. B. M. Curtis and W. A. Catterall, *J. Biol. Chem.* 258, 7280 (1983); *Biochemistry* 23, 2113 (1984). M. Borsotto *et al.*, *Biochem. Biophys. Res. Commun.* 122, 1357 (1984).
4. J. B. Gurdon *et al.*, *Nature (London)* 233, 177 (1971); C. B. Gunderlen *et al.*, *ibid.* 308, 421 (1984); K. M. Houamed *et al.*, *ibid.* 310, 318 (1984).
5. J. L. Maller, *Adv. Cyclic Nucleotide Res.* 15, 295 (1983); Y. Oron *et al.*, *Nature (London)* 313, 141 (1985); I. Lotan *et al.*, *Mol. Pharmacol.* 28, 170 (1985).
6. R. Miledi, *Proc. Soc. London Ser. B.* 215, 491 (1982); M. F. Barish, *J. Physiol. (London)* 342, 309 (1983).
7. T. Kazazoglou *et al.*, *FEBS Lett.* 164, 75 (1983).
8. K. Kusano *et al.*, *J. Physiol. (London)* 328, 143 (1982); C. Van Renterghem *et al.*, *Proc. R. Soc. London Ser. B.* 223, 389 (1985).
9. K. Kusano *et al.*, *Nature (London)* 270, 739 (1977); N. M. Dascal and E. M. Landau, *Life Sci.* 27, 1423 (1980); N. Dascal *et al.*, *J. Physiol. (London)* 352, 551 (1984).
10. B. P. Bean, *J. Gen. Physiol.* 86, 1 (1985); B. Nilius *et al.*, *Nature (London)* 316, 443 (1985).
11. J. N. Dumont, *J. Morphol.* 136, 153 (1972).
12. C. M. O'Connor and L. D. Smith, *Dev. Biol.* 52, 318 (1976).

13. We thank J. P. Leonard, A. M. Gurney, J. M. Nerbonne, J. Nargeot, and D. Fambrough for helpful discussion. N.D. was a Myron R. Rantrell Fellow, T.P.S. was a postdoctoral fellow of the American Heart Association/Greater Los Angeles Affiliate, and H.L. was a fellow of the Deutsche Forschungsgemeinschaft. Supported by NIH grants GM-10991 and GM-29836.

13 August 1985; accepted 17 December 1985

Sequence and Expression of Human Estrogen Receptor Complementary DNA

GEOFFREY L. GREENE, PAUL GILNA, MICHAEL WATERFIELD, ANDREW BAKER, YVONNE HORT, JOHN SHINE

The mechanism by which the estrogen receptor and other steroid hormone receptors regulate gene expression in eukaryotic cells is not well understood. In this study, a complementary DNA clone containing the entire translated portion of the messenger RNA for the estrogen receptor from MCF-7 human breast cancer cells was sequenced and then expressed in Chinese hamster ovary (CHO-K1) cells to give a functional protein. An open reading frame of 1785 nucleotides in the complementary DNA corresponded to a polypeptide of 595 amino acids and a molecular weight of 66,200, which is in good agreement with published molecular weight values of 65,000 to 70,000 for the estrogen receptor. Homogenates of transformed Chinese hamster ovary cells contained a protein that bound [^{3}H]estradiol and sedimented as a 4S complex in salt-containing sucrose gradients and as an 8 to 9S complex in the absence of salt. Interaction of this receptor-[^{3}H]estradiol complex with a monoclonal antibody that is specific for primate ER confirms the identity of the expressed complementary DNA as human estrogen receptor. Amino acid sequence comparisons revealed significant regional homology among the human estrogen receptor, the human glucocorticoid receptor, and the putative v-*erb*A oncogene product. This suggests that steroid receptor genes and the avian erythroblastosis viral oncogene are derived from a common primordial gene. The homologous region, which is rich in cysteine, lysine, and arginine, may represent the DNA-binding domain of these proteins.

THE REGULATION OF GENE EXPRESSION in eukaryotic cells by estrogens and other steriod hormones involves the interaction of specific intracellular receptor proteins with the genome, resulting in the activation of selected sets of responsive genes (*1*). As a consequence, DNA synthesis in certain target cells is altered, and there are changes in the synthesis of specific RNA's and proteins involved in the regulation of cell proliferation, differentiation, and physiologic function in diverse tissues. In addition, steriod hormones and their receptors appear to be involved in the regulation of abnormal growth in various tumors and tumor cell lines (*2*). Recent data from several laboratories (*3*) suggest that steroids may exert their effects by binding directly to an intranuclear receptor molecule that is weakly associated with nuclear components in the absence of ligand. Binding of hormone to its receptor results in conversion of the receptor-steroid complex to a form that associates with high affinity to one or more nuclear components. The molecular nature of this association and of the subsequent modulation of specific gene transcription is not known, although a number of nuclear acceptor sites have been proposed. These include specific DNA sequences (*4*), the nuclear matrix (*5*), and acidic nonhistone protein-DNA complexes (*6*). Although distinct steroid- and DNA-binding domains have been postulated to exist in all steroid receptors, few data are available on the detailed structure, composition, and chemical properties of the subunit that binds both steriod and DNA, and virtually nothing is known about the possible involvement of other components that do not bind steroid.

Determination of the primary structure of the estrogen receptor (ER) and expression of this molecule in homologous and heterologous systems can provide valuable information about structure-function relationships at a molecular level. Although ER is distributed in a tissue-specific manner, many of these cell types also express receptors for several other steroid hormones (*7*). Thus, it is likely that the specificity of control of responsive elements by steroids is determined, at least in part, by the primary structure of the receptor protein. Like other steroid receptors, hormone-occupied ER appears to recognize discrete DNA sequences that are generally upstream of transcriptional start sites in responsive genes. In the prolactin gene, footprinting analysis revealed a specific binding site for ER about 2 kb upstream of the start site (*8*); similar analyses of genes responsive to progestins and glucocorticoids revealed binding sites

G. L. Greene and P. Gilna, Ben May Laboratory for Cancer Research, University of Chicago, 5841 South Maryland Avenue, Chicago, IL 60637.
M. Waterfield, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, United Kingdom.
A. Baker, Y. Hort, J. Shine, California Biotechnology Inc., Palo Alto 94304.

located 150 to 200 nucleotides upstream of the transcriptional start sites (*9*). The corresponding binding elements on the receptor-hormone complexes have not yet been defined. However, it seems likely that functional domains with common elements will be found in the various steroid receptor proteins, especially if a discrete DNA-binding region exists.

The isolation of complementary DNA (cDNA) clones corresponding to part or all of the translated sequence of ER messenger RNA (mRNA) from MCF-7 human breast cancer cells has been reported (*10*). These sequences were identified in randomly primed λgt10 and λgt11 MCF-7 cDNA libraries by screening either with monoclonal ER antibodies or with synthetic oligonucleotides corresponding to two peptide sequences obtained from purified MCF-7 human ER. Among the cDNA clones isolated

```
                                                                    (-292)  GAATTCCAAAATTGTGATGTTTCTTGTATTTTTGATGAAGGAGAAATACTGT   -241
AATGATCACTGTTTACACTATGTACACTTTAGGCCAGCCCTTTGTAGGGTTATACAAACTCAAAGCACACCGCACCCGCAGGCTCCCGGGGCAGGGCCGGGGCCAGAGCTCCCGTGTCGC   -121
CGGGACATGCGCTGCGTCGCCCTCTAACCTCGGGCTGTGCTCTTTTTCCAGGTGGCCCGCCGGTTTCTGACGCCTTCTGCCCTGCGGGGACACGGTCTGCACCCTGCCCCGCGGCCACGGACC     -1

                    10                  20                  30                  40
MetThrMetThrLeuHisThrLysAlaSerGlyMetAlaLeuLeuHisGlnIleGlnGlyAsnGluLeuGluProLeuAsnArgProGlnLeuLysIleProLeuGluArgProLeuGly
ATGACCATGACCCTCCACACCAAAGCATCTGGGATGGCCCTACTGCATCAGATCCAAGGGAACGAGCTGGAGCCCCTGAACCGTCCGCAGCTCAAGATCCCCCTGGAGCGGCCCCTGGGC   120

                    50                  60                  70                  80
GluValTyrLeuAspSerSerLysProAlaValTyrAsnTyrProGluGlyAlaAlaTyrGluPheAsnAlaAlaAlaAlaAlaAsnAlaGlnValTyrGlyGlnThrGlyLeuProTyr
GAGGTGTACCTGGACAGCAGCAAGCCCGCCGTGTACAACTACCCCGAGGGCGCCGCCTACGAGTTCAACGCCGCGGCCGCCGCCAACGCGCAGGTCTACGGTCAGACCGGCCTCCCCTAC   240

                    90                  100                 110                 120
GlyProGlySerGluAlaAlaAlaPheGlySerAsnGlyLeuGlyGlyPheProProLeuAsnSerValSerProSerProLeuMetLeuLeuHisProProProGlnLeuSerProPhe
GGCCCCGGGTCTGAGGCTGCGGCGTTCGGCTCCAACGGCCTGGGGGGTTTCCCCCCACTCAACAGCGTGTCTCCGAGCCCGCTGATGCTACTGCACCCGCCGCCGCAGCTGTCGCCTTTC   360

                    130                 140                 150                 160
LeuGluProHisGlyGlnGlnValProTyrTyrLeuGluAsnGluProSerGlyTyrThrValArgGluAlaGlyProProAlaPheTyrArgProAsnSerAspAsnArgArgGlnGly
CTGCAGCCCCACGGCCAGCAGGTGCCCTACTACCTGGAGAACGAGCCCAGCGGCTACACGGTGCGCGAGGCCGGCCCGCCGCCGGCATTCTACAGGCCAAATTCAGATAATCGGACGCCAGGGT   480

                    170                 180                 190                 200
GlyArgGluArgLeuAlaSerThrAsnAspLysGlySerMetAlaMetGluSerAlaLysGluThrArgTyrCysAlaValCysAsnAspTyrAlaSerGlyTyrHisTyrGlyValTrp
GGCAGAGAAAGATTGGCCAGTACCAATGACAAGGGAAGTATGGCTATGGAATCTGCCAAGGAGACTCGCTACTGTGCAGTGTGCAATGACTATGCTTCAGGCTACCATTATGGAGTCTGG   600

                    210                 220                 230                 240
SerCysGluGlyCysLysAlaPhePheLysArgSerIleGlnGlyHisAsnAspTyrMetCysProAlaThrAsnGlnCysThrIleAspLysAsnArgArgLysSerCysGlnAlaCys
TCCTGTGAGGGCTGCAAGGCCTTCTTCAAGAGAAGTATTCAAGGACATAACGACTATATGTGTCCAGCCACCAACCAGTGCACCATTGATAAAAACAGGAGGAAGAGCTGCCAGGCCTGC   720

                    250                 260                 270                 280
ArgLeuArgLysCysTyrGluValGlyMetMetLysGlyGlyIleArgLysAspArgArgGlyGlyArgMetLeuLysHisLysArgGluArgAspAspGlyGluGlyArgGlyGluVal
CGGCTCCGCAAATGCTACGAAGTGGGAATGATGAAAGGTGGGATACGAAAAGACCGAAGAGGAGGGAGAATGTTGAAACACAAGCGCCAGAGAGATGATGGGGAGGGCAGGGGTGAAGTC   840

                    290                 300                 310                 320
GlySerAlaGlyAspMetArgAlaAlaAsnLeuTrpProSerProLeuMetIleLysArgSerLysLysAsnSerLeuAlaLeuSerLeuThrAlaAspGlnMetValSerAlaLeuLeu
GGGTCTGCTGGAGACATGAGAGCTGCCAACTTTGGCCAAGCCCCGCTCATGATCAAACGCTCTAAGAAGAACAGCCTGGCCTTGTCCCTGACGGCCGACCAGATGGTCAGTGCCTTGTTG   960

                    330                 340                 350                 360
AspAlaGluProProIleLeuTyrSerGluTyrAspProThrArgProPheSerGluAlaSerMetMetGlyLeuLeuThrAsnLeuAlaAspArgGluLeuValHisMetIleAsnTrp
GATGCTGAGCCCCCCATACTCTATTCCGAGTATGATCCTACCAGACCCTTCAGTGAAGCTTCGATGATGGGCTTACTGACCAACCTGGCAGACAGGGAGCTGGTTCACATGATCAACTGG  1080

                    370                 380                 390                 400
AlaLysArgValProGlyPheValAsnLeuThrLeuHisAspGlnValHisLeuLeuGluCysAlaTrpLeuGluIleLeuMetIleGlyLeuValTrpArgSerMetGluHisProVal
GCGAAGAGGGTGCCAGGCTTTGTGGATTTGACCCTCCATGATCAGGTCCACCTTCTAGAATGTGCCTGGCTAGAGATCCTGATGATTGGTCTCGTCTGGCGCTCCATGGAGCACCCAGTG  1200

                    410                 420                 430                 440
LysLeuLeuPheAlaProAsnLeuLeuLeuAspArgAsnGlnGlyLysCysValGluGlyMetValGluIlePheAspMetLeuLeuAlaThrSerSerArgPheArgMetMetAsnLeu
AAGCTACTGTTTGCTCCTAACTTGCTCTTGGACAGGAACCAGGGAAAATGTGTAGAGGGCATGGTGGAGATCTTCGACATGCTGCTGGCTACATCATCTCGGTTCCGCATGATGAATCTG  1320

                    450                 460                 470                 480
GlnGlyGluGluPheValCysLeuLysSerIleIleLeuLeuAsnSerGlyValTyrThrPheLeuSerSerThrLeuLysSerLeuGluGluLysAspHisIleHisArgValLeuAsp
CAGGGAGAGGAGTTTGTGTGCCTCAAATCTATTATTTTGCTTAATTCTGGAGTGTACACATTTCTGTCCAGCACCCTGAAGTCTCTGGAAGAGAAGGACCATATCCACCGAGTCCTGGAC  1440

                    490                 500                 510                 520
LysIleThrAspThrLeuIleHisLeuMetAlaLysAlaGlyLeuThrLeuGlnGlnGlnHisGlnArgLeuAlaGlnLeuLeuLeuIleLeuSerHisIleArgHisMetSerAsnLys
AAGATCACAGACACTTTGATCCACCTGATGGCCAAGGCAGGCCTGACCCTGCAGCAGCAGCACCAGCGGCTGGCCCAGCTCCTCCTCATCCTCTCCCACATCAGGCACATGAGTAACAAA  1560

                    530                 540                 550                 560
GlyMetGluHisLeuTyrSerMetLysCysLysAsnValValProLeuTyrAspLeuLeuLeuGluMetLeuAspAlaHisArgLeuHisAlaProThrSerArgGlyGlyAlaSerVal
GGCATGGAGCATCTGTACAGCATGAAGTGCAAGAACGTGGTGCCCCTCTATGACCTGCTGCTGGAGATGCTGGACGCCCACCGCCTACATGCCCCCACTAGCCGTGGAGGGGCATCCGTG  1680

                    570                 580                 590
GluGluThrAspGlnSerHisLeuAlaThrAlaGlySerThrSerSerHisSerLeuGlnLysTyrTyrIleThrGlyGluAlaGluGlyPheProAlaThrValSTOP
GAGGAGACGGACCAAAGCCACTTGGCCACTGCGGGCTCTACTTCATCGCATTCCTTGCAAAAGTATTACATCACGGGGGAGGCAGAGGGTTTCCCTGCCACAGTCTGAGAGCTCCCTGGC  1800
```

Fig. 1. OR8 cDNA and predicted amino acid sequence of human estrogen receptor from MCF-7 human breast cancer cells. The complete open reading frame consists of nucleotides 1 to 1785, out of a total of 2092. Termination codons (TGA) before and after the coding sequence are underlined. Peptide sequences obtained from purified MCF-7 ER after cleavage with cyanogen bromide or trypsin are also underlined. Numbers above the sequence refer to amino acid positions and numbers to the right indicate nucleotide locations. The unusual length of the 5' leader sequence upstream from the TGA codon at nucleotides −54 to −52 suggests that it may contain non-ER fusion sequence resulting from a cloning artifact. The sequencing strategy was as follows: OR8 cDNA was subcloned into the Eco RI site of M13mp9. Clones containing both orientations of the cDNA were isolated and subjected to DNA sequence analysis by the method of Sanger *et al.* (*13*), with a series of specific oligonucleotide primers homologous to either the adjacent M13 sequence or to previously determined cDNA sequences.

by oligonucleotide hybridization was a 2.1-kb cDNA (OR8) that cross-hybridized with all other cDNA's and that contained the expected sequences for the two ER peptides. In addition, this cDNA hybridized selectively to a 6.2-kb polyadenylated RNA which, when translated in vitro in the presence of [^{35}S]methionine, coded for the synthesis of immunoreactive 65-kD ER, as well as for a smaller amount of an immunoreactive 46-kD protein (*10*). The molecular size of the major product is similar to published values of 65 to 70 kD for ER from several sources (*11*). The identity of the smaller peptide is not known, but it may represent an in vitro degradation product of the 65-kD ER. The size of the mRNA for ER indicates that it is likely that a large portion is not translated. A 65-kD protein would require about 1.8 kb of coding sequence; this would leave more than 4 kb untranslated. This untranslated region is likely to be at the 3' end of the gene, as found in several other receptor mRNA's, including the human glucocorticoid receptor mRNA, which also contains a long 3'-untranslated region (*12*).

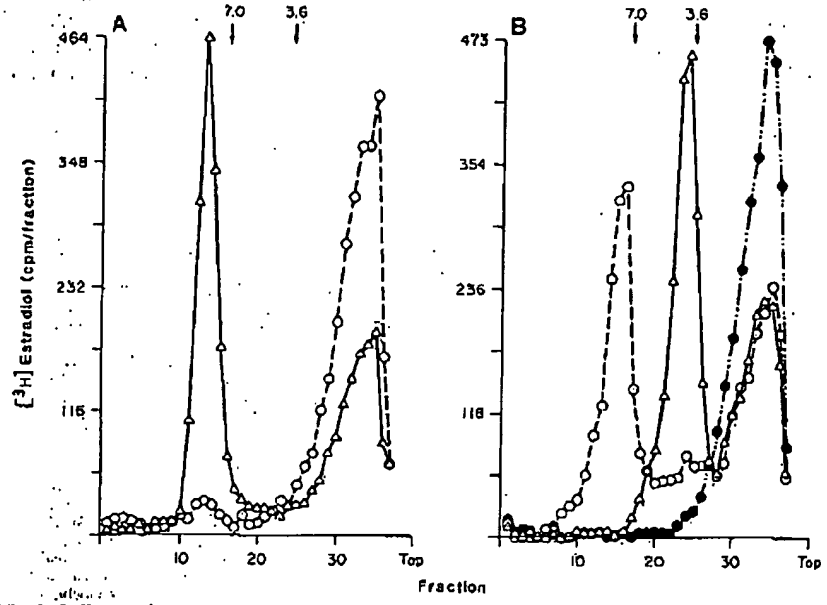

Fig. 2. Sedimentation analysis of human estrogen receptor expressed by OR8 cDNA in CHO-K1 cells and labeled with [^{3}H]estradiol. (A) Sedimentation profiles in low-salt gradients (10 m*M* KCl) of OR8-expressed human ER (Δ—Δ) or extract from untransformed CHO-K1 cells (●- -●) labeled with [^{3}H]estradiol (0.5 n*M*). (B) Sedimentation profiles in high-salt gradients (400 m*M* KCl) of expressed ER labeled with [^{3}H]estradiol (0.5 n*M*) and incubated with (○- -○) or without (Δ—Δ) D75P3γ immunoglobulin G (IgG), and expressed ER labeled with [^{3}H]estradiol (0.5 n*M*) plus unlabeled estradiol (100 n*M*) (●- · · -●). Untransformed CHO-K1 cells and cells transformed with the expression vector described below were grown in monolayer culture to confluency in a zinc-supplemented Dulbecco's minimal essential medium plus Ham's F-12 medium (*22*). After removal from the substrate with EDTA, the cells were homogenized by Polytron disruption in a buffer containing 10 m*M* tris (*p*H 7.4) and 20 m*M* sodium molybdate (*16*). Homogenates were centrifuged at 253,000*g* for 30 minutes and the supernatant fractions were labeled with 0.5 n*M* [^{3}H]estradiol (57 Ci/mmol), with or without a 200-fold molar excess of nonradioactive estradiol, for 60 minutes at 4°C. Experiments with excess diethylstilbestrol in place of estradiol showed the same pattern of radioactivity on the gradient. In a separate experiment, aliquots (200 μl) of labeled extract were incubated for 60 minutes at 4°C either in the presence or absence of rat monoclonal ER antibody (10 μg of D75P3γ, D547Spγ, or H222Spγ in a final volume of 220 μl). For sedimentation analyses, aliquots (200 μl) of labeled extract or incubation mixture were layered onto linear 10 to 30 percent sucrose gradients (3.5 ml), prepared in 10 m*M* tris, 10 m*M* sodium molybdate, 1.5 m*M* EDTA, *p*H 7.4, and either 10 m*M* KCl (low salt) or 400 m*M* KCl (high salt), and centrifuged at 0°C for 15 hours at 253,000*g*. Successive 100-μl fractions were collected and radioactivity was measured in Triton X-100 toluene scintillation mixture at 35 percent counting efficiency. [^{14}C]Ovalbumin (3.6*S*) and [^{14}C]IgG (7.0*S*) were used as sedimentation markers in parallel gradients; their positions are designated by arrows. The OR8 expression vector was constructed by subcloning the OR8 cDNA into pMTpn (*14*). This vector consists of a 870-bp Hind III–Bam HI fragment from the human metallothionein II gene (*14*) inserted into the polylinker region of pUC9. This genomic fragment contains metal regulatory regions, glucocorticoid-receptor binding sites, and the promoter, transcription start, and 5'-untranslated region of the metallothionein II gene. To provide for transcription stop and polyadenyl addition signals, the 600-kb Sma I–Eco RI fragment from the 3' end of the human growth hormone gene was placed on the 3' side of the OR8 cDNA. The vector also contains the SV40 origin of replication and enhancer sequences immediately 5' to the metallothionein promoter.

The 2.1-kb OR8 cDNA insert was the only cDNA long enough to contain the entire coding sequence for a 65-kD protein. Since this insert cross-hybridizes with all of the cloned cDNA's, regardless of selection method (*10*), OR8 should correspond to most or all of the MCF-7 human ER. To test this, the nucleotide sequence of this clone was determined by the method of Sanger *et al.* (*13*) according to the strategy described in Fig. 1. An open reading frame of 1785 nucleotides is the sequence encoding the human estrogen receptor, and corresponds to 595 amino acids and a calculated molecular weight of 66,200. The translation initiation site was assigned to the methionine codon at nucleotides 1 to 3 because this ATG triplet is the first to appear downstream from the in-frame terminator TGA at nucleotides −54 to −52. Although the actual initiation site has not been unequivocally established, amino acid sequence obtained from a cyanogen bromide fragment of MCF-7 ER corresponds to residues 12 to 26 (Fig. 1), which is very close to the proposed start site. The codon specifying the valine at position 595 is followed by a TGA translation termination codon. Peptide sequences from purified ER occur throughout the proposed open reading frame, including one that corresponds to residues 543 to 560 near the carboxyl terminus of the polypeptide (Fig. 1). Thus, all of the coding sequence for MCF-7 human ER is present in the OR8 cDNA insert.

To determine whether the OR8 cDNA would code for the synthesis of functional human ER in a heterologous cell system, OR8 was inserted into a pBR vector containing a metallothionein promoter and SV40 enhancer sequence (*14*), as described in Fig. 2. This OR8-containing plasmid was then used to transform Chinese hamster ovary cells (CHO-K1) (*15*), and cell homogenates were analyzed for the expression of human ER in a form capable of binding estradiol. Sedimentation analysis of a low-salt extract labeled with [^{3}H]estradiol revealed the presence of a receptor-[^{3}H]estradiol complex which sedimented at 8–9*S* in 10 m*M* KCl (Fig. 2A) and at 4*S* in 0.4*M* KCl (Fig. 2B). This complex reacted with three different monoclonal ER antibodies to form 8*S* immune complexes (Fig. 2B). One of these antibodies is specific for primate ER (D75P3γ) (*16, 17*). [^{3}H]Estradiol-binding in both the 4*S* and 8*S* complexes was abolished by the addition of a 200-fold molar excess of nonradioactive estradiol (Fig. 2B) or diethylstilbestrol. Extracts of untransformed CHO-K1 cells did not contain receptor-[^{3}H]estradiol complexes (Fig. 2A). In addition, the concentration of the receptor-[^{3}H]estradiol complex was more than

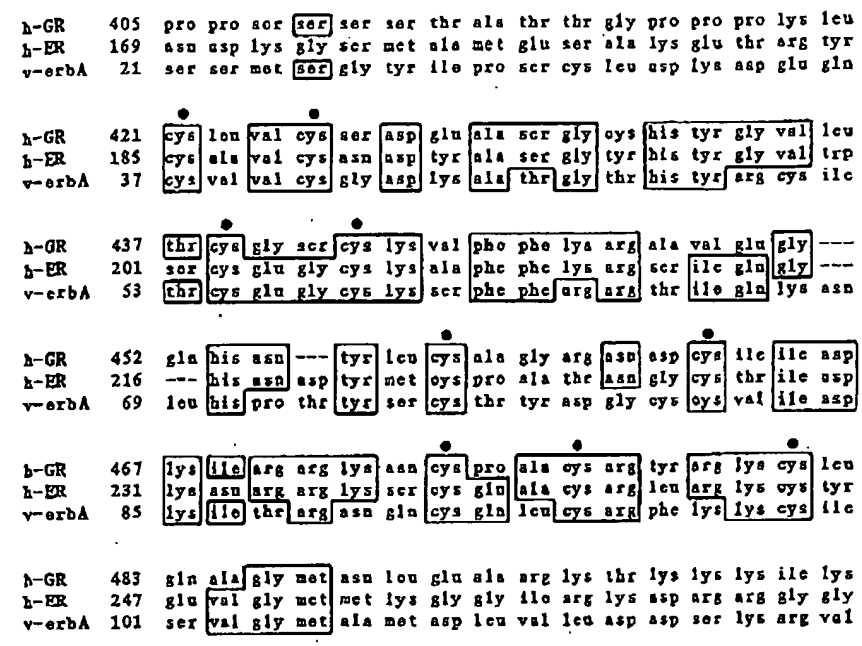

Fig. 3. Amino acid sequence alignment of the cysteine-, lysine-, and arginine-rich region of MCF-7 human ER, human GR, and putative v-*erb*A oncogene product. Amino acid residues 185 to 250 from ER were aligned with residues 421 to 486 from GR and residues 37 to 104 from $p75^{gag\text{-}erbA}$; common residues are boxed and gaps are indicated by dashes. Matching cysteine residues are indicated by dots above the sequence.

doubled by including $10^{-4}M$ Zn^{2+} in the culture medium for 24 hours prior to cell harvest. This result is consistent with the induction of the metallothionein promoter by zinc (*18*). The formation of 8–10*S* salt-sensitive receptor-hormone complexes in hypotonic extracts of responsive cells is a hallmark of steroid receptors, although the biological significance of these multimeric complexes has not been established. It is interesting that, although CHO-K1 cells appear to express little or no ER, the human ER expressed by OR8 cDNA in these cells forms an 8 to 9*S* complex when occupied by [^{3}H]estradiol under hypotonic conditions. This suggests either that this complex is a multimer of steroid-binding subunits or that associated nonsteroid-binding components are present in nontarget cells.

Identification of the locations and properties of the functional domains on the receptor protein can help establish the mechanism by which ER regulates gene transcription. Immunochemical analyses of partially proteolyzed MCF-7 human ER have identified at least two regions which are separable by enzymatic cleavage: a DNA-binding region and a steroid-binding region (*17*). Similar analyses of rat glucocorticoid receptor (GR) revealed a third "immunogenic" domain near the amino terminus of GR (*19, 20*) which is probably not present in ER. When the amino acid sequences of human ER and GR (*12, 20*) were compared, a striking homology was observed in a region rich in cysteine, lysine, and arginine (Fig. 3). This region occurs 300 to 350 amino acids from the carboxyl terminus in both ER and GR. Between residues 185 and 250 for ER and residues 421 and 486 for GR (66 amino acids) there are 40 matches (61 percent). A similar homology exists among ER, GR (*20*), and a region of the putative avian erythroblastosis virus oncogene protein $p75^{gag\text{-}erbA}$ (*20, 21*), from residue 37 to 104. All nine cysteines of ER are conserved in this region and nine out of ten GR and v-*erb*A cysteines are conserved. The abundance of cysteines in such a small region is unusual, and it has been suggested that this region and the associated basic amino acids may represent a DNA-binding domain of GR and v-*erb*A (*20*). Interestingly, this is the only region of significant sequence homology between human ER and GR, although both molecules contain a proline-rich region of unknown function located upstream from the region rich in cysteine, lysine, and arginine and a relatively hydrophobic region at the carboxyl terminus. On the basis of analysis of proteolytic and cDNA fragments of ER and GR (*10, 17, 19*) and of the nucleotide sequence of a form of GR that does not bind steroid (*20*), this hydrophobic region has been suggested to be the steroid-binding domain. If the primary amino acid structures of both proteins are aligned according to Fig. 3, much of the amino terminal immunogenic region present in GR is absent in ER, suggesting that this region may be important in distinguishing the action of GR from ER.

Thus, there is a strong relationship among two steroid receptors and the *erb*A proto-oncogenes, indicating that they are derived from a common primordial gene. Although the mechanisms by which steroid receptors modulate transcription and by which v-*erb*A promotes transformation are unknown, the common feature of these molecules is a domain that may be involved in DNA recognition. Further elucidation of the role of this domain and other functional regions in the regulation of gene expression should follow from in vitro and in vivo studies of the interaction of genetically altered receptor molecules with hormone-responsive genes.

REFERENCES AND NOTES

1. G. M. Ringold, *Annu. Rev. Pharmacol. Toxicol.* 25, 529 (1985); J. N. Anderson, in *Biological Regulation and Development*, R. P. Goldberger and K. R. Yamamoto, Eds. (Plenum, New York, 1985), vol. 3B, p. 169; B. S. Katzenellenbogen, *Annu. Rev. Physiol.* 42, 17 (1980); J. Gorski and F. Gannon, *ibid.* 38, 425 (1976); C. W. Bardin, L. P. Bullock, N. C. Mills, Y.-C. Lin, S. Jacobs, in *Receptors and Hormone Action*, B. W. O'Malley and L. Birnbaumer, Eds. (Academic Press, New York, 1978), vol. 2, p. 83; J. D. Baxter and P. H. Forsham, *Am. J. Med.* 53, 573 (1972); E. V. Jensen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 59, 632 (1968); K. R. Yamamoto and B. M. Alberts, *Annu. Rev. Biochem.* 45, 721 (1976); M. R. Sherman and J. Stevens, *Annu. Rev. Physiol.* 46, 83 (1984).
2. M. Lippman, in *Breast Cancer: Trends in Research and Treatment* (Raven, New York, 1975); M. Lippmann and G. Bolan, *Nature (London)* 256, 592 (1975).
3. P. J. Sheridan, J. M. Buchanan, V. C. Anselmo, P. M. Martin, *Nature (London)* 282, 579 (1979); W. V. Welshons, M. E. Lieberman, J. Gorski, *ibid.* 307, 747 (1984); W. J. King and G. L. Greene, *ibid.*, p. 745; M. F. Press, N. A. Nousek-Goebl, G. L. Greene, *J. Histochem. Cytochem.* 33, 915 (1985).
4. A. C. B. Cato, S. Geisse, H. M. Westphal, M. Beato, *EMBO J.* 3, 2771 (1984); J. P. Jost, M. Geisse, M. Seldran, *Proc. Natl. Acad. Sci. U.S.A.* 82, 988 (1985); J. G. Compton, W. T. Schrader, B. W. O'Malley, *ibid.* 80, 16 (1983); F. Payvar *et al.*, *Cell* 35, 381 (1983).
5. E. R. Barrack and D. S. Coffey, *J. Biol. Chem.* 255, 7265 (1980); *Endocrinology* 113, 430 (1983).
6. T. C. Spelsberg *et al.*, *Recent Prog. Horm. Res.* 39, 463 (1983).
7. K. B. Horwitz and W. L. McGuire, *J. Biol. Chem.* 253, 2223 (1978); R. D. Palmiter, P. B. Moore, E. R. Mulvihill, S. Emtage, *Cell* 8, 557 (1976).
8. E. Mulvihill, J. P. LePennec, P. Chambon, *Cell* 28, 621 (1982); F. Payvar *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 78, 6628 (1981); M. V. Govindan, E. Spiess, J. E. Majors, *ibid.* 79, 5157 (1982); D. Dean, R. Gope, B. J. Knoll, M. E. Riser, B. W. O'Malley, *J. Biol. Chem.* 259, 9967 (1984); D. Von der Ahe, S. Janich, C. Scheidereit, M. Beato, *Nature (London)* 313, 706 (1985); P. Chambon *et al.*, *Recent Prog. Horm. Res.* 40, 1 (1984).
9. R. A. Maurer, *DNA* 4, 1 (1985); M. G. Rosenfeld, personal communication.
10. P. Walter *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 82, 7889 (1985).
11. G. Rodeuilh *et al.*, *Eur. J. Biochem.* 106, 481 (1980); D. B. Lubahn, K. S. McCarty, Jr., K. S. McCarty, Sr., *J. Biol. Chem.* 260, 2215 (1985); J. A. Katzenellenbogen *et al.*, *ibid.* 258, 3487 (1983); D. Sakai and J. Gorski, *Endocrinology* 115, 2379 (1984); T. R. Van Oosbree, U. H. Kim, G. C. Mueller, *Anal. Biochem.* 136, 321 (1984).
12. S. M. Hollenberg *et al.*, *Nature (London)* 318, 635 (1985).
13. F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl.*

Acad. Sci. U.S.A. 74, 5463 (1977); P. H. Schreiner and R. Cortese, *J. Mol. Biol.* 129, 169 (1979).
14. M. Karin and R. I. Richards, *Nature (London)* 299, 797 (1982); L. C. Fritz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
15. T. T. Puck, *J. Exp. Med.* 108, 945 (1958).
16. G. L. Greene, C. Nolan, J. P. Engler, E. V. Jensen, *Proc. Natl. Acad. Sci. U.S.A.* 77, 5115 (1980).
17. G. L. Greene, N. B. Sobel, W. J. King, E. V. Jensen, *J. Steroid Biochem.* 20, 51 (1984).
18. R. L. Brinster, H. Y. Chen, R. Warren, A. Sarthy, R. D. Palmiter, *Nature (London)* 296, 39 (1982).
19. J. Carlstedt-Duke, S. Okret, O. Wrange, J.-A. Gustafsson, *Proc. Natl. Acad. Sci. U.S.A.* 79, 4260 (1982); O. Wrange, S. Okret, M. Radojcic, J. Carlstedt-Duke, J.-A. Gustafsson, *J. Biol. Chem.* 259, 4534 (1984); H. G. Dellweg, A. Hotz, K. Mugele, U. Gehring, *EMBO J.* 1, 285 (1982).
20. C. Weinberger, S. M. Hollenberg, M. G. Rosenfeld, R. M. Evans, *Nature (London)* 318, 670 (1985).
21. B. Debuire *et al.*, *Science* 224, 1456 (1984).
22. R. G. Ham, *Proc. Natl. Acad. Sci. U.S.A.* 53, 288 (1965).
23. We thank J. Miller for help with DNA sequencing, K. Harris for technical assistance, R. Myers-Knits for preparing the manuscript, and A. B. Zielinski for assistance and encouragement. This research was supported by research grants from the American Cancer Society (BC-86) and Abbott Laboratories, by the National Institutes of Health (CA-02897, and HD17703), by the National Science Foundation (INT-8313623), and by the Women's Board of the University of Chicago Cancer Research Foundation.

21 January 1986; accepted 31 January 1986

An Ancient Developmental Induction: Heat-Shock Proteins Induced in Sporulation and Oogenesis

STEPHEN KURTZ, JANICE ROSSI, LAWRENCE PETKO, SUSAN LINDQUIST

Every eukaryotic and prokaryotic organism tested to date synthesizes a small number of heat-shock proteins in response to heat and other forms of stress. A particular pattern of heat-shock gene expression was observed during ascospore development in *Saccharomyces*: heat-shock proteins hsp26 and hsp84 were strongly induced, whereas hsp70, the most highly conserved of these proteins, was neither induced nor inducible by heat shock. Instead, two proteins related to hsp70 were induced. A strikingly similar pattern of expression occurs during oogenesis in *Drosophila*, suggesting that it may be one of the earliest developmental pathways to evolve in eukaryotic cells.

ALL ORGANISMS RESPOND TO MILD elevations in temperature by coordinately synthesizing a small set of heat-shock proteins. The exact number of proteins induced varies in different organisms, but in all cases proteins of approximately 84 and 70 kilodaltons (hsp84 and hsp70) are among the most prominent species. These proteins have been highly conserved in evolution. The hsp70 proteins of *Drosophila* and yeast have 72 percent amino acid identity (*1*) and their respective hsp84 proteins have 63 percent identity (*2, 3*). Most organisms also produce heat-shock proteins of 20 to 30 kD. *Drosophila* cells produce four closely related proteins of 28, 26, 23, and 22 kD. Cells of the yeast *Saccharomyces cerevisiae* produce only one prominent small protein with a molecular mass of 26 kD. These small heat-shock proteins have not been conserved to the same extent as hsp70 and hsp84, but nucleic acid sequence analysis has demonstrated homology among the proteins of insects, vertebrates, and nematodes (*4–5*). Furthermore, the small heat-shock proteins of *Drosophila*, yeast, and tomatoes form particles of highly conserved morphology (*6, 7*).

Although the specific functions of the heat-shock proteins are not yet known, some of them are expressed during oogenesis and pupation in *Drosophila* (*8–12*), suggesting that they play a role in normal development as well as in the response to stress. To investigate developmental regulation of the heat-shock genes in the yeast *S. cerevisiae*, we examined sporulating cells. Since these cells do not efficiently take up radiolabeled amino acids, gene expression was determined with DNA probes and antibodies. In the experiment represented in Fig. 1, diploid cells of the strain AP3 reached the tetranucleate stage 8 to 10 hours after transfer to nitrogen-deficient medium and sporulation was complete at 24 hours. Total cellular RNA's were isolated at various times during sporulation, electrophoretically separated, and analyzed by hybridization with cloned probes for the heat-shock genes.

Messenger RNA (mRNA) for hsp26 was induced early in sporulation, eventually reaching a concentration higher than that achieved during a 1-hour heat shock (Fig. 1a). Messenger RNA for hsp84 was also induced during sporulation (Fig. 1b). The timing of its accumulation was different from that of the hsp26 message. The maximum level of induction was comparable to that achieved with a 1-hour heat shock.

The hsp70 gene family in *Saccharomyces* contains two different classes of heat-inducible genes encoding 70-kD proteins. Transcripts from one class, represented here by clone YG100, are observed at low levels at 25°C and at much higher levels at 36°C. Transcripts of the other, represented by clone YG107, are observed only at temperatures above 38°C (*13*). Neither class was induced during sporulation (Fig. 1c). Moreover, as can be seen with longer exposures, as sporulation proceeded, the concentration of the YG100 message dropped below the basal level observed during normal vegetative growth.

We examined the expression of heat-shock RNA's in several *S. cerevisiae* strains of widely divergent genotypes. Messages for hsp26 and hsp84 are induced at different times in strains that sporulate at different rates, but they are always induced strongly. Neither class of hsp70 message was induced during sporulation in any strain. Thus, unlike the coordinate induction of these genes during heat shock, their induction during development is uncoupled; only a particular subset of heat-shock genes is induced.

This pattern of heat-shock gene expression is remarkably similar to one reported to occur during normal oogenesis in *Drosophila*. In adult females, RNA's for hsp26, hsp28, and hsp84 are induced in ovarian nurse cells and passed into the developing oocyte (*11*). As with meiosis in *S. cerevisiae*, this developmental induction differs from heat-shock induction in that mRNA for hsp70 does not accumulate. In fact, in late egg chambers and early embryos, hsp70 is not induced even with heat shock (*11*). This is significant, since, in virtually all other tissues, hsp70 is the protein most strongly induced by heat.

To determine whether hsp70 is heat-inducible during sporulation, we removed portions of a sporulating culture at various times during development and subjected them to heat shock at 39°C. RNA's from these cells were hybridized with probes for the two classes of hsp70 genes. Both were inducible during the early stages of sporulation; neither was inducible during the final stages of spore maturation (Fig. 2a). This change apparently occurred before general transcriptional inactivation of the spore genome, since an mRNA encoding a 21.5-kD sporulation-specific polypeptide accumulated after hsp70 became refractory to induction (Fig. 2b).

We translated RNA's from sporulating cells in vitro and found that the heat-shock messages they contained were fully translatable. However, the fact that these RNA's can be translated in vitro provides no infor-

Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, IL 60637.